Tracking People Across Disjoint Camera Views

A Thesis Submitted For the Degree of Doctorate of Philosophy

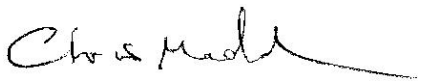
By

Christopher Madden

Faculty of Information Technology
University of Technology, Sydney
Australia

July 3, 2009

I, **Christopher Madden**, certify that the work in this thesis titled "**Tracking People Across Disjoint Camera Views**" has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text. I also certify that the thesis has been written by myself. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. The undersigned certify that they have read this thesis and that in their opinions it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.



Christopher Madden

Date: July 3, 2009

Principal Supervisor: Massimo Piccardi

Acknowledgements

I would like to thank all those people who have been involved in some way with the creation of this thesis. My biggest thanks goes to Professor Massimo Piccardi for being my supervisor and assisting me with learning all of the areas I have needed to know for this work. And for filling me in on the areas that I still haven't mastered yet. To my loving wife and family thank you for all the caring support. I would like to promise that I will talk a little less about video surveillance from now on; however we both know that is unlikely to happen. I would also like to thank my friends and colleagues who have assisted me throughout my candidature. Especially those who have read, or even reread, this thesis and helped me to improve it to its current state. I would also like to thank you, the reader, for your time as this thesis is not only written to explore my knowledge of the area, but hopefully to pass some of that on to you.

Abstract

Tracking people around surveillance systems is becoming increasingly important in the current security conscious environment. This thesis presents a framework to automatically track the movements of individual people in large video camera networks, even where there are gaps between camera views. It is designed to assist security operators, or police investigations by providing additional information about the location of individuals throughout the surveillance area. Footage from an existing surveillance system has been used to test the framework under real conditions. The framework uses the similarity of robust shape and appearance features to match tracks. These features are extracted to build an object feature model as people move within a single camera view, which can be compared across cameras. The integration of matching similarities in the temporal domain increases the robustness to errors of many kinds. Frames with significant segmentation errors can be automatically detected and removed based upon their lack of similarity to the other models within the same track, increasing robustness.

The shape and appearance features used to generate the object models are based upon features humans habitually use for identifying individuals. They include a height estimate, a Major Colour Representation (MCR) of the individuals global colours, and estimates of the colours of the upper and lower portions of clothing. The fusion of these features is shown to be complementary, providing increased discrimination between individuals. The MCR colour features are improved through the mitigation of illumination changes using controlled equalisation, which improves the accuracy in matching colour under normal surveillance conditions and requires no training or scene knowledge. The incorporation of other features into this framework is also relatively straightforward.

This track matching framework was tested upon four individuals across two video cameras of an existing surveillance system. Existing infrastructure and actors were used to ensure that ground truth is available. Specific cases were constructed to test the limitations of the system when similar clothing is worn. In the data, the height difference ranges from 5 to 30 centimetres, and individuals may only be wearing 50% of similar clothing colours. The accuracy of matching an individual was as high as 91% with only 5% false alarms when all the system components were used. This may not become a fully automated system, but could be used in semi-automated or human assisted systems, or as the basis for further research into improved automated surveillance. Application areas range from forensic surveillance to the matching of the movements of key individuals throughout a surveillance network and possibly even target location.

Contents

List of Figures

List of Tables